Consonance in Music and Mathematics: Application to Temperaments and Orchestration

Constança Martins de Castro Simas
constanca.simas@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

December 2014

Abstract

Since ancient Greece there is evidence of the relation between Music and Mathematics. This article aims to be one more contribution to the study of the links connecting both fields. Two main topics will be investigated and discussed along the following sections. Although, both of them will relate to two important concepts in music: *consonance* and *timbre*. One of the approaches relates with the tuning of scales performed in ancient times, and the other with acoustics and consonance in the orchestra nowadays. Pythagoras studied the relation between rational numbers and pure sounds, tuning scales in a way that would preserve perfect consonances. However, this turned out to have an irregular consonance in the different intervals constituting the scale. The goal of what is called a temperament is to find a tuning system capable of optimizing these features within a scale. Therefore, a computational method shall be developed to output the consonance between two musical notes, in the range $[0, 1]$. The second matter developed along this article deals with the different sounds in the orchestra. The "colors" of the instruments in the orchestra are analysed through a mathematical procedure. A sound produced by an instrument is defined by its harmonic spectrum, representing its timbre. To derive these spectra, we shall use concepts from Fourier theory, specially the Discrete Fourier Transform. This process is initiated with the recording of instrument sounds, and then completed with the analysis of the soundwaves.
Keywords: consonance, harmonic spectrum, Fourier transform

1. Introduction

Musicians understand certain concepts like consonance and timbre by using their auditory sensitivity and memory. This paper also intends to formalize mathematically some of this intuitive notions. The areas in mathematics that will mainly be used are:

- Number theory, used to represent numerical intervals in a musical scale;
- Algebraic notions and vector spaces for the calculation of consonance;
- Fourier theory for the representation of the frequency spectrum of a sound;
- Differential equations to represent the wave equation that produce a sound.

Number theory is applied in the sense that the relation between different notes in a scale can be represented by numbers, more specifically, fractions. This is a way of defining numerically a note in a scale, but we can also describe it as a vector. A vector of a sound contains the resonant frequencies that constitute it. It may also contain the weights of the intensity correspondent to each resonant frequency.

The first two areas mentioned above shall be useful to calculate consonance ahead. The last two, will be used for the mathematical analysis of the instruments in the orchestra. After recording the sounds we can find an equation defining the soundwave and discover the harmonic spectrum, by using the Fourier transform.

Although covering many subjects, the main idea of this thesis is to combine them in a common goal: an attempt to demystify some rumours in music, using solid mathematical concepts.

2. Background

Along this section we shall introduce some music definitions and also briefly explain the major concepts of Fourier theory.

2.1. Sound and Consonance

Young people usually learn music by singing the notes of a scale and intuitively corresponding to each note one sound. Although, when it is said that an instrument sounds at a pitch of $f\,Hz$, that sound is essentially periodic with frequency f. According to *Fourier theory*, a sound is decomposed into a sum of sine and cosine waves at integer multiples of the frequency f. We call to the component of the sound with frequency f the *fundamental* and the components with frequencies $m \times f$, $m \in \mathbf{N}$, the mth *harmonics*. The components of a sound can be sometimes inharmonic and in that case they are called *overtones* of the fundamental. A sound is also characterized by the following features:

- **Pitch**, corresponds to the frequency in Hz of the fundamental;
- **Timbre**, which defines the quality of the sound of a certain instrument;
- **Intensity**, defining if a sound is more or less loud;
- **Duration**, which can be measured in seconds or through rhythm.

The concept of timbre is what distinguishes one instrument's sound from another and this is a consequence of two different issues. One is the number of resonant harmonics, and the other is the intensity of each of them. It is common for the intensity to generally decay when reaching high harmonics, but the shape of this effect varies from instrument to instrument, serving almost as a fingerprint. The *Harmonic series* or *spectrum* of a sound corresponds to the harmonics representing it. Pythagoras wanted to take this notion one step forward and so he tried to use two sounds at the same time.

Definition 1 (Interval)**.** An interval is a combination of two notes at different or equal pitch. It can be represented by the ratio between their frequencies.

Definition 2 (Octave)**.** An octave is an interval such that its ratio is $2^n/1$, $n \in \mathbf{N}$, and the notes played are the same with the difference that one is higher in pitch. By the expression *reducing intervals to an octave* we mean that we divide an interval ratio by two until it belongs to the interval $[1, 2]$.

It makes sense to assume that if we use the harmonics of a note as fundamentals to play a second note, then we obtain an interval which is represented by a ratio of integer numbers. This process allows us to find every possible interval within an octave (Table 1). When hearing an interval we realise the superposition of two harmonic spectra and that gives us the concept of *consonance*.

Definition 3 (Consonance)**.** Consonance exists when two notes in different pitch are played at the same time and sound pleasant together. This happens when these two notes have overtones of the respective fundamental in common. Also, the interval is more consonant if the superposed harmonics are the ones that resound more, that is, the ones of smaller order [3].

Unison	1/1
Minor Tone	10/9
Major Tone	9/8
Minor Third	6/5
Major Third	5/4
Fourth	4/3
Fifth	3/2
Minor Sixth	8/5
Major Sixth	5/3
Octave	2/1

Table 1: Ratios of the most important intervals

2.2. Wave equation and Fourier analysis

Musical instruments are mechanic-acoustic systems since they are constituted by two types of vibrations. The ones in a solid object, the instrument itself, and the propagation in a fluid, which is the air in the acoustic point of view [4]. These vibrating movements are periodic oscillations yielding a simple harmonic motion described by

$$F = -kx = m\ddot{x} \Longrightarrow \ddot{x}m + kx = 0 \Leftrightarrow \ddot{x} + \frac{k}{m}x = 0, \tag{1}$$

where k is a constant, m the mass of the particles and x the distance from the equilibrium position. The sinusoidal movement is described by $x(t) = A\sin(\omega t + \phi)$ where ω is the *angular velocity* or *frequency of the movement*, in $rad\,s^{-1}$, and ϕ is the *initial phase*, in rad. So we we must check the conditions for which this is a solution of equation (1).

$$\frac{\partial x}{\partial t} = -A\omega\sin(\omega t + \phi),$$

replacing on the equation we get:

$$-mA\omega\sin(\omega t + \phi) + kA\sin(\omega t + \phi) = 0 \Leftrightarrow$$
$$\Leftrightarrow (-m\omega + k)A\sin(\omega t + \phi) = 0.$$

Considering the points where $A\sin(\omega t + \phi) \neq 0$ then $-m\omega + k = 0$. Therefore we obtain the relation $\omega = \sqrt{\frac{k}{m}}$ or $f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$. For the complex sound there is an alternative representation given by the *Fourier series*, which essentially results on the sum of the components of different frequencies in a wave.

$$F(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n\cos(2n\pi ft) + b_n\sin(2n\pi ft)), \tag{2}$$

$$a_m = 2f \int_0^T \cos(2m\pi ft)F(t)dt, \quad m > 0$$
$$b_m = 2f \int_0^T \sin(2m\pi ft)F(t)dt, \quad m > 0.$$

This representation is possible only if F is a periodic function with period T, continuous and having a bounded continuous derivative except in a finite number of points in $[0, T]$. In this case, the series as defined above converges to F at all points where it is continuous.

Now that we know how to approximate a soundwave by a Fourier series, we also want a method that allows us to get information about its frequency spectrum. This can be obtained by using the Fourier transform which converts signals from a time domain to a frequency domain. The Fourier transform is given by:

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-2\pi i\omega t}dt. \quad (3)$$

In order to apply the transform, f must be an integrable function with real domain. By integrable function we mean that $\int |f|d\mu < +\infty$. The behaviour of the transform is characterized by the Riemann-Lebesgue lemma, which states that the Fourier transform of an integrable function tends to zero when the frequency tends to infinity, that is, $\lim_{\omega\to\infty} \int_{-\infty}^{+\infty} f(t)e^{-2\pi i\omega t}dt = 0$.

The Fourier transform defined above is used for a continuous infinite domain. Although, when the available data is a wave in sound format, it is necessary to work with a discrete domain. To obtain the discrete version of the transform, first we imagine the continuous soundwave, f. When sampling the wave we perform an assumption of discrete time so that $f(t) \to f(t_k) = f_k$ and $t_k = k\Delta$, where Δ is the gap size between two values in time. If we have a list of N samples then $k = 0, ..., N-1$ and the Discrete Fourier Transform is defined as follows:

$$\mathcal{F}_n = \sum_{k=0}^{N-1} f_k e^{\frac{-2\pi ink}{N}}. \quad (4)$$

The output of this transformation is a complex number containing information on the amplitude and phase of the sinusoid, in a frequency domain. For the purpose of calculating the frequency spectrum of a sound we only need the amplitude of each component and so we shall only consider $|\mathcal{F}_n|$.

3. Sound recording and mathematical implementation

This section of the article is dedicated to report how the instruments were recorded and how were the functions created to analyse them. First we explain the routines implemented to calculate frequency spectra and then the program constructed to obtain the relative consonance of two notes.

3.1. Frequency spectra of instruments in an orchestra

The sounds analysed were all recorded in the anechoic chamber of *Instituto Superior Técnico*. This type of chamber disables reflections and insulates noises from the outside. The instruments were recorded with a condenser microphone and an external soundcard was used to control the whole set. The resultant sounds of this process are in WAV format, which is an uncompressed sound format. This is used to obtain the best representation of the amplitude relation in the recording. Also, the resultant sound is sampled with a rate of $441000\,Hz$, whose meaning will be explained later in the article. The recordings were performed with original orchestra instruments and by music students of *Academia Nacional Superior de Orquestra*.

In order to obtain the frequency spectra, using the *Mathematica* platform, we must understand how the relevant data of a wave is contained in a sound of WAV format. The *Sampling Rate* is the number of equally spaced amplitude samples kept in a list for one second of a signal. The choice of the rate is based on the following theorem.

Theorem 1 (Nyquist-Shannon sampling theorem)**.** If a function f contains no frequencies higher than $B\,Hz$, it is completely determined by giving its ordinates at a series of points spaced $\frac{1}{2B}$ seconds apart [6].

Since the human hearing range goes from $20\,Hz$ to $20\,kHz$ it makes sense that the sampling rate is near the double of this interval, therefore the regular use of $44100\,Hz$.

The list of discrete values obtained through the WAV format can be used to apply the Discrete Fourier Transform. Although, it would be a little imprecise to apply it in a really small part of the wave. So we use *Welch's method* which consists in splitting a part of the wave into D overlapping segments and applying the transform to each of them [8]. The average of the results for all the segments is the one taken to draw the frequency spectrum. To obtain the spectrum it's also required to transform the list of values into a list of points in which we consider a domain of frequencies. Let *data* be the list of amplitudes in a frequency domain, L the length of the list and *rate* the sampling rate. If the gap between samples in the time domain is $\frac{1}{rate}$ then in the frequency domain the gap will be $\frac{rate}{L}$. Therefore, the coordinates in the frequency domain correspondent to the amplitudes in *data* are $\frac{n\,rate}{L}$, $n \in 0, ..., L-1$.

The final version of the spectrum, however, consists in a domain of natural numbers n representing the nth harmonics and a relative amplitude from 0 to 1. This can be obtained by dividing the frequencies on the domain by the fundamental and the amplitudes by the maximum of the plot. Finally, it is possible to interpolate the results to draw a final function consisting in weights from 0 to 1 for each of the harmonics (see Figure 1).

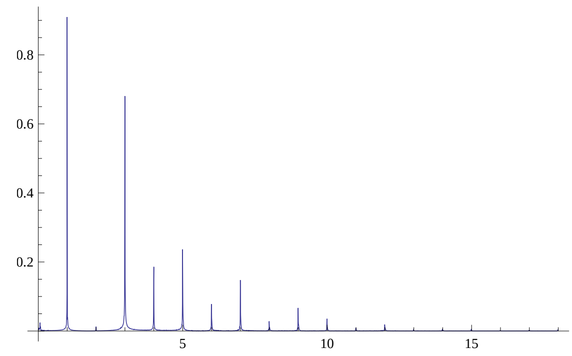

Figure 1: Plot of the Frequency Spectrum

3.2. A program to compute consonance

There are three features that must be considered in order to obtain a reliable consonance between two notes. First of all we must have the notes in a list format, in this case, listing the frequencies of the corresponding harmonics. Then, it's possible to group the harmonics of equal frequency between the lists of two notes and count how many are in common. If these are many, by the definition of consonance, we get higher consonance. Although, as seen before with the analysis of harmonic spectra, the different harmonics of one note have different intensities. Therefore, we consider a function giving a weight between 0 and 1 to each of the harmonics, just like the one drawn in Figure 1.

Let's suppose that we are searching for the consonance of the interval consisting on two notes A and B with different fundamental frequencies f_1 and g_1. We also have a weight function for each note, w and p for notes A and B, respectively. These weight functions relate the ith harmonic with its relative amplitude $w(i) := w_i$ and $p(i) := p_i$. The notes A and B are represented by $\{f_1, f_2, ..., f_n\}$ and $\{g_1, g_2, ..., g_n\}$ where f_i and g_i are the harmonics, $i \in \{1, 2, ..., n\}$. Let $\{(f_{l_1}, g_{j_1}), ..., (f_{l_m}, g_{j_m})\}, l, j \in \{1, 2, ..., n\}$ be the list of the pairs of frequencies in common, where $f_{l_1} = g_{j_1}, ..., f_{l_m} = g_{j_m}$. Using this, the consonance of the interval between notes A and B is given by:

$$c(A,B) = \frac{\parallel (w_{l_1}p_{j_1}, ..., w_{l_m}p_{j_m}) \parallel}{\parallel (w_1 p_1, ..., w_n p_n) \parallel}. \tag{5}$$

It makes sense to consider the euclidean norm of the vector $(w_{l_1}p_{j_1}, ..., w_{l_m}p_{j_m})$ since the multiplication of two weights belonging in $[0, 1]$ is still a weight in that interval and maintains the relativity between smaller and bigger weights. Then, the norm considered above is divided by the same vector specific for the unison played in instruments with any weight functions. This procedure allows for the maximum consonance to be obtained only with a unison in equal or different timbres, and all the remaining values of consonance belong to $[0, 1[$.

There is still one last thing to add to the calculation of consonance and that is the concept of *Critical Bandwidth*. For example, if we input frequencies 101 Hz and 200 Hz in the program, it will output a consonance very close to 0. This happens because the overtones of the two fundamental frequencies do not overlap at all. The problem is that human ear can't distinguish these small gaps of frequencies and it would find no difference between the interval using a 100 Hz pitch and the other one with 101 Hz. So, each time we hear a sound at a frequency f and then vary the pitch on a certain interval in hertz, our ear can't distinguish if we're still hearing the same sound or not. That interval is the critical bandwidth of the original frequency f. The size of the critical bandwidth depends on the fundamental frequency of a note according to $CB = 94 + 71F^{\frac{3}{2}}$, where CB is the critical bandwidth in Hz and F the central frequency in KHz [7]. When a note's frequency withdraws from the central frequency, the dissonance between them increases. A model of this decay of consonance between close frequencies was carried out by *Plomp* and *Levelt* who worked on an experimental analysis of consonance and dissonance [5]. The equation obtained was:

$$\begin{cases} 1 - 4|x|e^{1-4|x|}, \ |x| < \frac{1}{4} \\ 0, \ |x| \geq \frac{1}{4}. \end{cases} \tag{6}$$

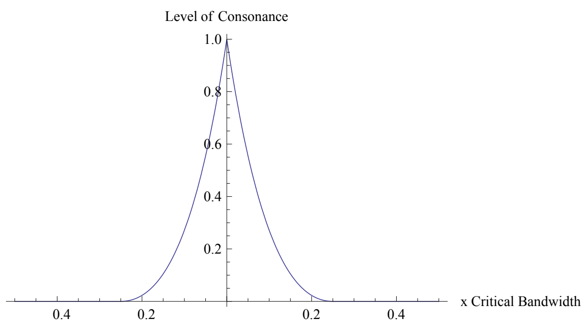

Figure 2: Plomp and Levelt's results for consonance on a fraction of the critical bandwidth

On the xx axis we have the central frequency of a note when $x = 0$ and a total dissonance is obtained when the difference of pitch reaches a quarter of the critical bandwidth of the central frequency.

To implement this feature we consider again the lists of harmonics of both notes A and B and their respective weights: $\{(f_1, w_1), ..., (f_n, w_n)\}$ and $\{(g_1, p_1), ..., (g_n, p_n)\}$ where f_i and g_i are the frequencies of the harmonics and w_i, p_i the respective weights, $i \in \{1, ..., n\}$. Firstly, we need to combine all the frequency elements in one list with the ones in a second list. Suppose that the program is comparing (f_i, w_i) with (g_k, p_k), where $i, k \in \{1, ..., n\}$. To f_i and g_k we apply the function $CB = 94+71F^{\frac{3}{2}}$ in order to discover the critical bandwidth of both frequencies, which we shall call CBf and CBg, for f_i and g_k respectively. The next step is to apply Plomp and Levelt's function 6 to $\frac{f_i - g_k}{CBf}$ and to $\frac{f_i - g_k}{CBg}$ to obtain a relative consonance of f_i when belonging to the interval $[g_k - CBg, g_k + CBg]$ and the same for g_k when belonging to $[f_i - CBf, f_i + CBf]$. Let c_{ik} be the average value between these values of relative consonance.

After comparing all the combinations of harmonics, we can obtain a vector of the form:

$$v(A, B) = (w_1 p_1 c_{11}, w_1 p_2 c_{12}, ..., w_2 p_1 c_{21}, ... \\ ..., w_n p_1 c_{n1}, ..., w_n p_n c_{nn}).$$

Finally, we have the consonance between notes A and B, taking into account the critical bandwidth:

$$c(A, B) = \frac{\parallel v(A, B) \parallel}{\parallel (w_1 p_1, ..., w_n p_n) \parallel}, \tag{7}$$

This adapted version of consonance works similarly to the formula (5). The difference is that in this case we consider the additional weight of the critical bandwidth in the vector v. The normalization is performed by dividing the euclidean norm of v by the norm of the unison vector. It doesn't make any sense to consider the critical bandwidth for the case of the unison since $f_i = g_i$ where $i = 1, ..., n$.

4. Temperaments

The octave interval is traditionally divided into 12 equal intervals to form a *chromatic scale*. To tune a scale it would be common sense to use the acoustically pure intervals seen in Table 1. However, any method used to construct a scale always reaches an impure interval, as we will see next.

A way to tune all the notes is to add twelve consecutive fifths and see if it is possible to reach seven octaves:

$$\frac{12\ fifths}{7\ octaves} = \frac{(3/2)^{12}}{2^7} = \frac{3^{12}}{2^{19}} = \frac{531441}{524288} = \\ = 1,013643265.$$

The value above represents the *Pythagorean comma* and we conclude that adding intervals of fifth leads to a spiral and not a circle [2].

Adding four consecutive fifths and two octaves and a pure major third also leads to a gap:

$$\frac{4\ fifths}{2\ octaves\ +\ 1\ third} = \frac{(3/2)^4}{2^2 \times (5/4)} = \frac{3^4}{2^4 \times 5} = \\ = \frac{81}{80} = 1,0125.$$

This is called the *syntonic comma*. Any scale tuned in this manner would lead to a really big interval of third.

Finally, adding three consecutive pure major thirds doesn't correspond to one pure octave, and the difference is called a *diesis*:

$$\frac{1\ octave}{3\ major\ thirds} = \frac{2}{(5/4)^3} = \frac{2^7}{5^3} = \frac{128}{125} = 1,024.$$

A *temperament* is a way of compromising some of the pure intervals in a scale, in order to obtain the rigorous condition of the pure octave between the first and last notes. Along the following subsection, we present some of the most common temperaments designed over time [1].

4.1. Popular temperaments

One of the first solutions for the "comma problem" was the **Pythagorean Tuning**. The method used was to tune a sequence of fifths, passing by all the twelve notes of the scale. After tuning twelve fifths we reach the problem of the Pythagorean comma and that is why the last fifth has to be tuned narrower than the others, by one Pythagorean comma. For this reason, the intervals of third on this tuning are one syntonic comma wider than a pure one, which sounds almost out of tune. Therefore, the music written in ancient Greece uses mainly the intervals of fifth and octave, avoiding the thirds. The representation of the intervals with this temperament is in Table 2.

Another solution is the **Just Intonation** which consists in tuning a scale with intervals of small ratios between the beginning note and the following ones. All the ratios for the just scale can be obtained by listing the ratios of the harmonics of the fundamental of the scale. It is possible to find all these ratios by analysing only the first 30 harmonics of a note. The general problem associated to this kind of temperament is that if an instrument is tuned in a just major scale starting on C, the just major triad sounds very well but a chord on any other key can sound really harsh. So it is obvious that this complicates any type of modulation to different keys.

Unison	1
Minor Tone	256/243
Major Tone	9/8
Minor Third	32/27
Major Third	81/64
Fourth	4/3
Augmented Fourth	1024/729
Fifth	3/2
Minor Sixth	128/81
Major Sixth	27/16
Minor Seventh	16/9
Major Seventh	243/128
Octave	2

Table 2: Ratios of the intervals in a Pythagorean chromatic scale

The most common tunings used until the nineteenth century were the **Meantone Temperaments**. These consist in tuning a scale by making a cycle of fifths adjusted by a fraction of the syntonic comma. This allows for the thirds to be acoustically pure and the fifths still acceptable. One of the ways to tune it is to subtract a quarter of the comma to all the fifths except one. That last fifth, the *wolf fifth* turns out to be $\frac{7}{4}$ of a comma wider, in order to compensate the adjustments performed to the other fifths. Other way to organize this is to reduce $\frac{1}{4}$ of the comma to 8 fifths and add to other 4 which softens the effect of the wolf fifth. The same type of procedure can be applied with $\frac{1}{6}$ of the comma.

Finally we get to the temperament used nowadays, the **Equal Temperament**. It consists on distributing, in equal parts, the syntonic comma by the circle of fifths. Adding intervals corresponds to multiplying their ratios. Therefore, the ratios of this temperament consist in multiplying $2^{1/12}$ until we reach 2, the octave. This is the same as adding second minors of the same size until we have a chromatic scale

4.2. Application of the consonance program

Now, having the background of the existing temperaments in music, it is possible to use the program elaborated on chapter 3 to find out which temperaments sound better in each interval of a chromatic scale. To obtain an accurate comparison, all the temperaments are tested under the same conditions, that is, all the ratios are analysed with the same weight function. For that purpose we created a function interpolating an average of the harmonic peaks for every recorded instrument playing the central C (Figure 3).

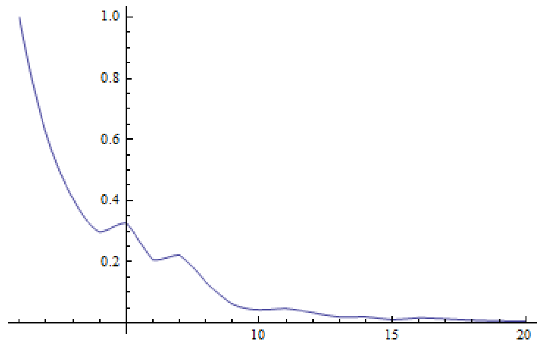

Figure 3: Interpolation of the average spectrum of all the instruments

The ratios of each temperament were taken as input for the consonance program, giving the results on Table 3.

	Equal	Just	1/6-Comma
Unis	1	1	1
2m	0,0044	0,0005	0,0210
2M	0,0088	0,0106	0,0075
3m	0,0310	0,0617	0,0349
3M	0,0472	0,0888	0,0647
4P	0,1045	0,1136	0,0973
4A	0,0304	0,0479	0,0479
5P	0,2193	0,2378	0,2046
6m	0,0227	0,0397	0,0139
6M	0,0587	0,1221	0,0750
7m	0,0153	0,0188	0,0133
7M	0,0018	0,0029	0,0025
8P	0,5961	0,5961	0,5961

Table 3: Consonance results for three temperaments: Equal, Just and 1/6-comma Meantone.

Observing the results, one can notice some expected values for the temperaments. For example, the equal temperament's major thirds are the most dissonant of all. In the just scale, the consonances for the intervals are higher than in all the other temperaments, except the minor second, augmented forth and major seven. These intervals are the ones considered to sound worse in music, so we can conclude that the just scale is the one showing the greatest contrasts between the consonances and dissonances.

As for the intervals in general, the fifths are all very consonant even though they are only pure in the just scales. The octaves, fifths and fourths are the most consonant, by this order, as it would be expected. The augmented fourths consonance is not as low as it might seem, historically the devil's interval. The values of consonance can be very low sometimes, since the spectrum of the instruments decays rapidly.

5. Timbres and Consonance in the Orchestra

The orchestra is a musical phenomenon which consists in a large ensemble of musicians playing instruments. The orchestra is divided in the following sections: strings, woodwinds, brass and percussion. Up to modern times, each of the orchestra sections increased a lot in size and this not only created an enormous amount of sound but also provided more timbral effects to the composer. The organization of the playing instruments in a music piece is called *orchestration*.

In order to study the consonance inside an orchestra, we must first analyse each instrument in terms of its frequency spectrum because it defines the timbre. This implies also checking the different ways of producing sound in each orchestra section and describe analytically an approximation of its characteristic soundwaves.

5.1. Strings

The vibration of strings amplified by the resonance box of the instrument is the core of the sound production of these instruments. Considering a string held at both ends, the vibrations are mainly transversal. The motion is represented by a variable y corresponding to the vertical displacement, x representing the position along the string and t, time. Suppose we have a segment of a string displaced as follows:

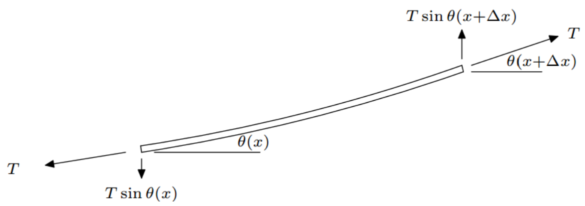

Figure 4: Representation of a string with an applied tension [2].

We consider x and Δx, two positions measured along the horizontal axis. $\theta(x)$ is the angle of the string, so $tan\theta(x) = \frac{\partial y}{\partial x}$. Since the string is held at both ends, this means that there is a tension applied in both sides of the string. We write T for the tension in Newton ($kg\,m/s^2$) and ρ for the linear density of the string, in kg/m.

In Figure 4 are represented the vertical components caused by the applied tension. For small angles, which is the case of transversal vibrations, $\tan\theta(x) \simeq \sin\theta(x)$. Now we want to check the vertical displacement:

$$\begin{aligned} T\tan\theta(x+\Delta x) - T\tan\theta(x) &= \\ = T\left(\frac{\partial y(x+\Delta x)}{\partial x} - \frac{\partial y}{\partial x}\right) &= \\ = T\Delta x \frac{\frac{\partial y(x+\Delta x)}{\partial x} - \frac{\partial y}{\partial x}}{\Delta x} \simeq T\Delta x \frac{\partial^2 y}{\partial x^2}. \end{aligned} \tag{8}$$

Since $m = \rho\Delta x$, where m is the mass of the string, and the acceleration in the vertical axis is $\frac{\partial^2 y}{\partial t^2}$, then we have that

$$\rho\Delta x \frac{\partial^2 y}{\partial t^2} \simeq T\Delta x \frac{\partial^2 y}{\partial x^2}.$$

So, we obtain the general equation to describe the motion of a vibrating string:

$$\frac{\partial^2 y}{\partial t^2} = c\frac{\partial^2 y}{\partial x^2}, \tag{9}$$

where $c = \sqrt{\frac{T}{\rho}}$.

The harmonic spectra of the different string instruments share most of the characteristics. Firstly, the sound of an open string has always more and louder harmonics than one pressed by a finger. Besides a little muffling caused by the placement of the finger, there is also the fact of reducing the string's length. The smaller the length, the smaller the wavelength of the harmonics. This implies that the vibrations are more constricted, leading to less resounding harmonics.

Generally, the more intense harmonics in the spectrum of a string instrument are the ones corresponding to octaves of the fundamental frequency and octaves of the open strings of the instrument. A string vibrating by sympathy, caused by the resonance of a sound, is a common effect and justifies the characteristic of the spectrum mentioned before. It is common to find peaks also in odd harmonics like the 7th and the 13th. All these characteristics can be observed in the violin spectrum in Figure 5.

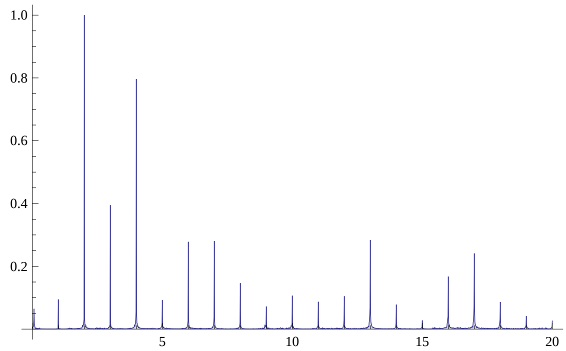

Figure 5: Harmonic spectrum of violin's open string G

There is a difference between the spectra of the instruments related with the size of the resonance box. The violin's box is proportional to the range of pitches it has to resonate, the viola and the cello would be impractical if they maintained the same proportion to the respective registers. This makes the violin the instrument with more resonant harmonics. The cello and viola only have similar spectra in the higher strings.

5.2. Woodwinds

The wave equation of a wind instrument is expressed by the vibrating air particles inside the tube. To calculate this we must consider two new variables, *air pressure* and *displacement*. Air pressure is usually at the ambient value ρ. When the air is compressed by the moving particles we obtain a pressure $P(x,t)$. The *acoustic pressure* is defined as $p(x,t) = P(x,t) - \rho$. The displacement is the change from a rest position in the tube and is defined by $\xi(x,t)$, the displacement of the air at position x and time t.

According to Hooke's law,

$$p = -B\frac{\partial \xi}{\partial x}, \tag{10}$$

where B is the *bulk modulus* of the air. Since $m = \rho\partial A dx$, where ∂A is the area of the surface where the air passes, then $F = -\partial p \partial A = \rho \partial A \partial x a$. So, we can define acceleration as $a = -\frac{1}{\rho}\frac{\partial p}{\partial x}$ and,

$$\frac{\partial p}{\partial x} = -\rho\frac{\partial \xi}{\partial t}. \tag{11}$$

Using equations (10) and (11) we obtain the following equations

$$\frac{\partial \xi}{\partial x} = \frac{1}{c}\frac{\partial \xi}{\partial t} \qquad \frac{\partial p}{\partial x} = \frac{1}{c}\frac{\partial p}{\partial t} \tag{12}$$

with $c = \sqrt{\frac{B}{\rho}}$. These are the equations of displacement and acoustic pressure, respectively.

We say that we have a *pressure node* where there is no fluctuation in the pressure, and an *antinode* when the change in the air pressure is maximum. A displacement node corresponds to a pressure antinode, since the particles movement is constrained by the high acoustic pressure, and vice versa.

The boundary conditions of equations (12) are dependent on the tube being open or closed. In the open tube the pressure nodes are in the extremities of the tube, since they are at the ambient air pressure. Figure 6 represents the displacement nodes and antinodes in an open tube.

In the case of a closed tube, like the clarinet, we have a pressure node in the sole open extremity and an antinode in the closed end. For that reason the modes only allow frequencies of the odd harmonics. The oboe and bassoon behave as open tubes of the same length. The conical bore of these instruments makes them work as an open tube squashed at the closed end.

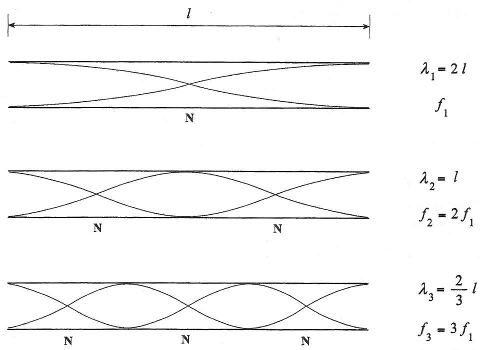

Figure 6: The first three modes of displacement nodes and antinodes for an open tube. N stands for node [4].

The woodwinds emit sounds in different pitches by covering and uncovering holes in the tube. This procedure has the effect of "shortening" the length since the air leaves the tube also by the holes. Another technique for reaching different pitches is to change the register. This can be obtained by exciting the vibration of the other modes of the tube, and not only the fundamental. The flute changes register by changing the air pressure blown into the tube, and the other reed instruments by a key that makes the reed vibrate in different speeds.

The woodwinds, unlike strings, have spectra with a more linear decay instead of harmonic peaks. The intensity of the harmonics generally grows until a peak, which varies depending on the register of the notes, and then decreases. In low sounds the peak usually starts around the 5th harmonic and as they go to upper registers the order of the peak harmonic decreases, until it reaches the fundamental.

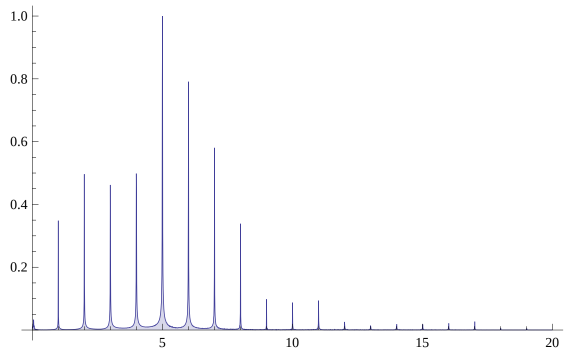

Figure 7: Harmonic spectra the note C4 in the oboe.

The woodwinds have more sound impact than the strings but the sound is not so rich, meaning that it doesn't have so many harmonics. One of the richer sounds of the oboe is represented in Figure 7 and the order of the harmonics ends essentially around the 11th harmonic while the strings usually have resonant harmonics until the 20th.

5.3. Brass

The propagation of sound in the brass instruments works as in the woodwinds, however the means to reproduce different notes are completelly different. The brass instruments contain a mouthpiece, allowing the lips to vibrate. The tension in the lips produces a change in the vibration frequency, which induces a harmonic to vibrate along the tube. Therefore, through this method the brass instruments can reproduce a harmonic sequence. The valves and slides enable the reproduction of chromatic notes. These mechanisms introduce additional tubing into the instruments, changing the tube length and thus, the whole fundamental for the harmonic series.

The lip makes the brass instrument work as a closed tube since there is reflection at the very small lips aperture. However, the tuba and the french horn have a conical bore so they work as an open tube, just like the oboe. The trombone and the trumpet have cylindrical pipes but the odd harmonics are manipulated into a whole harmonic series. The mouthpiece lowers the high pitch frequencies towards forming a harmonic series and the bell, on the contrary, works as a resonator of high frequencies, bringing the lower frequencies up. The bell is also the reason why brass instruments have high harmonics more intense than woodwinds. The artificial construction of the harmonic series may cause some lag between the higher overtones and the natural harmonics. This doesn't happen in the instruments of conical bore because they already reduce the "closed tube effect". Figure 8 shows that the brass harmonic spectra work in a similar way to the woodwinds, but there are more harmonics and generally more intense.

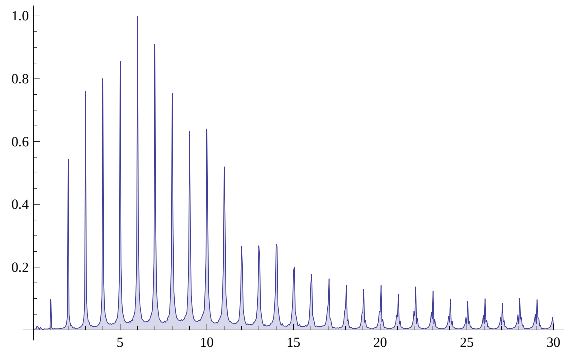

Figure 8: Harmonic spectra the note F1 in the tuba.

When the bell raises the lower resonances, none of them reaches the fundamental of the series. However, this non-existent fundamental can be produced in some brass instruments by vibrating the lips at its frequency because the higher harmonics completing the series provide a make-believe resonance, the so called *pedal tone*.

5.4. Percussion

The vibrations in a percussive membrane are transversal, and work as a three dimension string wave. The tension T in the membrane is uniform and the mass density ρ is now measured in kg/m^2, because we are considering area units. We picture a rectangle in the middle of a circular membrane to analyse the tension in the various axis, applied to the rectangle (Figure 9).

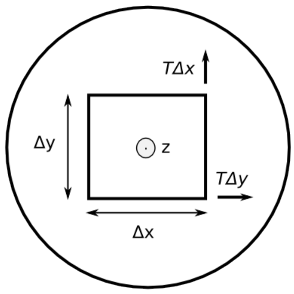

Figure 9: The vertical and horizontal tensions in the membrane.

Taking into account equation 8, we have that the difference in vertical components of the force (in the z axis) caused by the tension on the left to the right sides and the front to the back are respectively:

$$(T\Delta y)(\Delta x\frac{\partial^2 z}{\partial x}) \qquad (T\Delta x)(\Delta y\frac{\partial^2 z}{\partial y}). \tag{13}$$

So, the total vertical force is $T\Delta x\Delta y(\frac{\partial^2 z}{\partial x}+\frac{\partial^2 z}{\partial y})$. Since the mass of the rectangle is $\rho\Delta x\Delta y$, by Newton's law we have

$$T\Delta x\Delta y(\frac{\partial^2 z}{\partial x}+\frac{\partial^2 z}{\partial y}) \simeq (\rho\Delta x\Delta y)\frac{\partial^2 z}{\partial t} \Leftrightarrow$$
$$\Leftrightarrow c(\frac{\partial^2 z}{\partial x}+\frac{\partial^2 z}{\partial y}) = \frac{\partial^2 z}{\partial t}, \tag{14}$$

where $c=\sqrt{\frac{T}{\rho}}$ as usual. Transforming into polar coordinates we have the final equation,

$$\frac{\partial^2 z}{\partial t} = c(\frac{\partial^2 z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 z}{\partial \theta}+\frac{1}{r}\frac{\partial z}{\partial r}). \tag{15}$$

The solution of this equation uses *Bessel functions* which causes the frequencies of the harmonics to depend on their zeros. This leads to inharmonic overtones of the membrane's spectrum. For that reason, the common percussion instrument used in orchestra, the kettledrum, despite having a tuning system based on stretching and releasing the membrane always resounds a little out of tune.

5.5. Clarinet and oboe, the most consonant dissonance

To end this analysis on timbre we study the consonance of a small interval in the clarinet and the oboe. This case is peculiar because the characteristics of each timbre are totally distinct.

The following analysis is only a small part of what can be obtained with the consonance program. This subsection intends to find an answer to the following question: "How can we create the worst second minor between the oboe and the clarinet?". Note that the same question could be applied to any combination of instruments and intervals, and this means we can find the worst dissonances but also the best consonances. The program of consonance allows us to find the consonance for each interval in different or equal timbres. For this specific case we looked at the consonance of the minor second between C4 and C♯4. The following table displays the results of this interval played by oboes and clarinets in all possible combinations.

$obC4 + obC\sharp4$	$0,0715$
$obC4 + clC\sharp4$	$0,0928$
$clC4 + obC\sharp4$	$0,0983$
$clC4 + clC\sharp4$	$0,0879$

Table 4: Consonances of a minor second in the clarinet and the oboe

This results allow curious conclusions like, for example, when the interval is played by a clarinet and an oboe, the most consonant option is with the low note in the clarinet. This fact is perfectly acceptable since the clarinet, having more harmonics in these low notes, is probable to have more harmonics in common with the higher oboe sound. The second and last conclusion is that two oboes or two clarinets playing this interval sound more dissonant than one clarinet and one oboe. We can therefore say that this dissonance is reinforced by the clash of two equal timbres.

This exemplifies the type of analysis that one can do in specific cases of orchestration and composition. Wherefore, different excerpts of orchestral music can be analysed by using the consonance program.

6. Conclusions

We can draw different conclusions on the two ways of applying the notion of consonance, in temperaments and orchestration.

For the case of the temperaments, we conclude that although the equal temperament has been almost exclusive in music since the 19th century, it might not be the best choice. Also, before equal temperament, different keys had different "colors" and now they are a mere transposition of tone. The Meantone temperament, for example, is versatile in modulations from one scale to another allowing a different sonority for each of them, and still preserving an acceptable consonance of the intervals. A new perspective on the reuse of temperaments may allow a whole new universe of possibilities to show human emotion through music.

As for the second part, the idea of a program outputting which timbre options are better for a specific effect can be an important tool for a composer. Of course this is analysed only in terms of consonance, but that already allows the creation of different timbre environments.

The analysis of the spectrum for each note of an instrument is already of great importance to orchestration. Lets consider a typical example used in music composition: the use of an instrument with many intense harmonics to enrich the sound of other instrument. The trumpet playing in a high register emits a powerful sound but very thin and that is why the clarinet is sometimes used in unison.

The two applications of the program serve two different ambitions. One is to show that the ancient knowledge in temperaments was used to create effects nowadays lost. The other is to prove that a mathematical approach to the notion of timbre and consonance can assist in many ways to the composition of music.

References

[1] P. Asselin. *Musique et Temperament.* Jobert, 2000.

[2] D. J. Benson. *Music: A Mathematical Offering.* Cambridge University Press, 2007.

[3] H. Helmholtz. *On the Sensations of Tone.* Dover Publications, fourth edition, 1954.

[4] L. L. Henrique. *Acústica Musical.* Fundação Calouste Gulbenkian, 4th edition, 2011.

[5] R. Plomp and W. J. M. Levelt. Tonal consonance and critical bandwidth. *J. Acoust. Soc. Amer.*, 38, 1965.

[6] C. E. Shannon. Communications in the presence of noise. *Proc. IRE*, 37:10–21, January 1949.

[7] J. O. Smith and J. S. Abel. The bark and erb bilinear transforms. *IEEE Transactions on Speech and Audio Processing*, December 1999.

[8] P. D. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio Electroacoustics*, AU-15:70–73, 1967.